A data-driven approach to urban digitalization

**Keywords:** data, digitalization, analytics, interoperability, enterprise architecture

# Introduction

The Milan Municipality’s digital plan is inspired by the strategic goals of the Council and of the National Agency for the digitalization of Public Administration (AGID).

## The Council’s strategic goals

The strategic goals include technical innovation, transparency, public involvement, development, liveability and sustainability.

## A vision for IT

The strategic vision for IT that supports the above activities is based on the following principles:

* Information sharing in “open data” mode
* “Mobile first” approach
* Safe, scalable and harmonized IT architectures
* Well-defined application interfaces
* Full compliance with cyber-security standards
* Adoption of enabling platforms

# Methods

The above principles are enacted through 4 broad lines of action:

* Reorganizing the ICT Division
* Defining and implementing an Enterprise Architecture (EA) Model
* Making legacy systems interoperable and compliant with the EA
* Implementing innovation projects

## Reorganization

Some new divisions have been created in the IT Division. They oversee:

* Data management and integration (includes the statistical office)
* Digital lead and digital agenda
* Information security and technological innovation

## Enterprise Architecture

In the new EA, responsibilities have been distinctly divided between the IT directorate and the rest of the municipal divisions. The IT sector oversees the technological and application aspects of the IT systems, while the functional and business aspects fall under the responsibility of the other divisions.

## Tackling legacy systems

To implement all this, the city management support systems must comply with the EA and the interoperability principles. This is being achieved through an ongoing change process in which the systems are classified into macro-categories representing the different aspects of city management, and their data assets and interoperability are assessed. The resulting change process for each system is then ranked by priority.

## Innovation projects

The envisaged innovation projects include a system interoperability platform; use of artificial intelligence to improve user experience and system efficiency; an open data platform and integrated statistical system; a data lake with master data management; and a comprehensive geo-intelligence portal, to be used for both data dissemination and operational purposes.

# Results

The “new IT deal” at Municipality of Milan started in late 2017. The main results achieved to date in the interoperability, data management and integration areas are as follows:

* The open data portal is in place, and currently contains more than 430 datasets and a special section publishing graphical representations of the municipal budget;
* The geoportal has implemented about 20 new applications;
* The interoperability platform is in place, and its API store contains 20 interfaces to operational systems and experimental platforms;
* The data lake is in place, and data ingestion has been completed for several systems and is ongoing for several more;
* A set of indicators on the international economic placement of Milan has been published in the Osservatorio Milano website (<https://www.osservatoriomilanoscoreboard.it/>, in Italian);
* Some operational dashboards have been developed to help improve internal processes in the customer relations and human resource projects;
* A “big data” project has helped to identify weak points in the fine levying process.

# Conclusions

The following figure depicts the data governance architecture.



Figure 1. Data governance

Data from the various sources flow towards the data lake through the APIs. The data lake is divided into a “deep lake”, where data arrive and are enriched by operational metadata in a semi-automatic way, and a “smart lake”, where they will be integrated according to a model which is currently being defined.

The data-centred projects currently either use a single data source or integrate data sources on a case by case basis.